DOCUMENT RESUME

ED 262 107 TM 850 629

AUTHOR Hunt, Earl
TITLE The Design of Ballantines.
INSTITUTION Washington Univ., Seattle. Dept. of Psychology.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE Aug 85
CONTRACT N00014-84-K-5553
NOTE 26p.; Small print throughout document.
PUB TYPE Reports - Research/Technical (143) -- Computer Programs (101)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Algorithms; Analysis of Covariance; *Computer Graphics; *Correlation; Educational Research
IDENTIFIERS *Ballantines; PASCAL Programing Language

ABSTRACT

The widespread availability of computational graphics for personal computers has greatly increased the potential for visual display of research data. The correlation between two variables can be represented graphically by using circles to represent the variance associated with each circle, and by letting the overlap between the circles be proportional to the squared product-moment correlation between the variables. A procedure is given for locating the circles so that this relation is fulfilled. The method is extended to the three-variable case. The extended case can be used to construct graphical representations called ballantines. The PASCAL computer program for the ballantine procedures is appended. (Author/BS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED262107

THE DESIGN OF BALLANTINES

EARL HUNT

AUGUST 1985

This research was sponsored by:

Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research
Under Contract No. N00014-84-K-5553
Contract Authority No. NR 667-528

Approved for public release; distribution unlimited.

Reproduction in whole or in part is permitted for any purpose for the U.S. Government.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ERIC

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report 8 -	2. GOVT ACCESSION NO	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Design of Ballantines		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Earl Hunt		8. CONTRACT OR GRANT NUMBER(s) N00014-84-K-5553
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology, University of Washington, Seattle, Washington 98195		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs, Office of Naval Research, Arlington, Virginia 22217		12. REPORT DATE August 15, 1985
		13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS (of this report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Statistics, graphics, correlation, graphical displays, part correlation, partial correlation.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The correlation between two variables can be represented graphically by using circles to represent the variance associated with each circle, and letting the overlap between the circles be proportional to the squared product-moment correlation between the variables. A procedure is given for locating the circles so that this relation is fulfilled. The method is extended to the three variable case. The extended case can be used to construct graphical representations of part and partial correlations.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102 LF 014 6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

The Design of Ballantines

Earl Hunt

The University of Washington

Address correspondence to:

Earl Hunt

Psychology Department

University of Washington

Seattle, WA 98195

ERIC

Abstract

Shared variance can be expressed graphically by overlapping circles. A procedure is presented for locating the circles so that the graphical and statistical relations correspond exactly. The procedure is extended to represent part and partial correlations between three variables.

The widespread availability of computational graphics for personal computers has greatly increased the potential for visual displays of data. The display of pairwise correlations between two and three variables is of special interest to psychologists. To motivate the subsequent development, consider a case that arose in our own laboratory. College students participated in three tasks, an auditory dichotic listening task, a visual scanning task, and an arithmetic task. The correlations between the tasks were

(auditory, visual) = .42

(auditory, arithmentic) = .47

(visual, arithmetic) = .30

Our interest was in the extent to which variance was shared between pairs of tasks, with some portion of the variance in the third task "held constant". Part and partial correlations may be used to express the statistical relations. However this method of summarization was not appropriate for verbal presentations of our results, especially to audiences who were not familiar with advanced methods of correlational analysis.

An alternative to the statistical summary is to use a visual display, in which the variance of each variable is represented by a circle. Shared variance is represented by the overlap between two circles. If three variables are represented the resulting figure is called a <u>ballantine.</u> Several authors have advocated their use to represent covariation in three variable problems (e.g. Cohen and Cohen, 1975). The ballantine is a useful display of shared and unique variance because each component of variance can be identified visually in the geometric form. This can be seen in Figure 1, which is a ballantine representation of our data. The various part and partial correlations can be expressed in terms of the regions of overlap (a_1, a_2, a_3, and a_4) shown in the figure.

Figure 1 here

Obviously, ballantines are generated from simpler "two circle" figures that represent the variance-covariance relations between two variables, X and Y. This is shown in Figure 2. If representations such as Figures 1 and 2 are to portray data accurately the

preparation of a circle's area lying in the intersection region (Region A in Figure 2) should be exactly equal to r^2, the squared correlation between the appropriate variables. In fact, the ballantine of Figure 1 does fulfill this condition for our data. Figure 2 exactly represents the correlation between the auditory and visual detection measures. The purpose of this note is to explain how such figures may be constructed.

Figure 2 here

<u>The Underlying Geometric Relations.</u>

Let the circles X, Y, and Z stand for the variances of three variables, x, y, and z. Let the circles have a constant radius, R. This "visually standardizes" the variables by representing Var (x), Var (y) and Var (z) by circles with area πR^2. Two circles X, Y are said to be placed correctly with respect to each other if and only if the overlapping area contains the proportion of each circle equal to the squared correlation coefficient. In the case of Figure 1, area A is equal to

$$(1)\quad A = r^2_{xy}\pi R^2$$

The area of intersection of circles X and Y, both of radius R, is determined by the length of line L between the center of circle X (C_x) and the center of the circle Y (C_y), i.e. by L_{xy}/R. This is shown in Figure 3. Therefore, for fixed C_x, C_y may be located anywhere on the circle of radius L_{xy} centered on C_x. If we adopt the conventions that L_{xy} be horizontal and that X always lies to the left of Y, the locus of circle Y is thus determined once X is located and L_{xy} is determined.

Figure 3 here

The position of the third circle of a ballantine can be determined in a similar way. The center of circle Z (representing the variance of variable z) must lie on the circumference of a circle of radius L_{xz}, centered on C_x and on the circumference of a circle of radius L_{yz} centered on C_y. Since two non-identical circles intersect at either two or no points, there are two possible ballantines when the three variables share common variance. In one of these, circle Z lies above the line L_{xy}, in the other it lies below it. Either figure wuld be an

appropriate ballantine. Here circle Z will always lie below the horizontal. These relations are shown in Figure 4.

For the sake of completion two degenerate cases must be mentioned. If λ_{xy} = 1, then circles X and Z are identical (L_{xy} = 0), and similarly for X and Z and Y and Z. If λ_{xy} = 0, then L $\geq$2R, so that circles X and Y do not overlap. By convention the relation L_{xy} = 2R will be used, so that the circles for variables that do not share common variance will lie next to each other without overlapping.

Figure 4

<u>Trigonometric Relations</u>

An algorithm for determining the length of L_{xy} will now be presented. The identical algorithm, with a change of variable names, applies to L_{xz} and L_{yz}. Developing the algorithm is basically an exercise in high school trigonometry.

The algorithm will be described by referring to the lines and angles shown in Figure 3. Consider the segment bound by line AB and arc h. This has area 1/2 A, where A is the area of overlap. The value of A is defined by

(2) $1/2\ A = 1/2\ R^2(\alpha - \sin(\alpha))$ (Burington, 1948).

where α is measured by radians.

For a 'standard' circle, with R=1, equation (1) may be substituted into (2). Then, simplifying,

(3) $\lambda^2_{xy}\pi = \alpha - \sin(\alpha)$.

Note that if λ^2_{xy} is 1 α has the value of π (in radians). At this point the two circles will be identical. At the other extreme, if λ^2_{xy} is zero $\alpha = 0$. This establishes limits on α.

Equation (3) defines α implicitly, as a transcendental function of λ^2_{xy}. The value of α for a given value of λ^2_{xy} may be approximated to any desired degree of accuracy. The existance of a unique solution is ensured by the fact that the quantity $(\alpha - \sin(\alpha))$ increases monotonically from $,\pi$ throughout the range

of α. (The first derivative, $1 - \cos(\alpha)$, is non-negative for $0 \leq \alpha \leq \pi$). Once α is found, the value of R can be calculated directly. By inspection of Figure 3,

(4) $L_{xy} = 2R - 2h$.

However

(5) $R - h = R \cdot (\cos(\alpha/2))$

Substituting , and letting R = 1 to establish a scale,

(6) $L_{xy} = 2(\cos(\alpha/2))$.

Therefore the problem is solved if α can be determined. This can be done by finding the value of α that satisfies (3).

Computation

The computation of α for a given λ^2_{xy} is generally not feasible without a computer. Appendix 1 is a PASCAL program that executes the appropriate algorithm. It computes circle positions given the correlations for a two or three variable problem. The heart of the program is the procedure CONVERGE. For any value of λ^2, converge calculates α by successive approximations until α is

ERIC

within .0001 radians of its true value. The value of L is then computed by using equation (6).

It would be tedious to recompute the relations for every new case of a bivariate relation. Table 1 presents values of L_{xy}/R for and r^2_{xy} ranging from .00 to 1.00 in steps of .01. If a ballentine is to be drawn by hand Table 1 can be used to determine the radii of the circles to be used in the construction.

If the ballantines are to be drawn by computer graphics, let C_X be located at point (X_X, Y_X) in a Cartesian Co-ordinate system. A convenient position for $C_y(X_y, Y_y)$ is

$$(7)\quad X_y = X_X + L_{xy}$$
$$Y_y = Y_X$$

Locating C_z is slightly more complex. As Figure 4 shows, the three points C_X, C_y, and C_z define a triangle with sides L_{xy}, L_{xz}, and L_{yz}. Let θ be the interior angle of the triangle $\Delta\ C_X C_y C_z$ at point C_X. For ease of notation, let

$$(8)\quad a = L_{yz}$$
$$b = L_{xz}$$

$$c = L_{xy}$$
$$s = 1/2\ (a+b+c)$$

and

$$(9)\quad V = ((s-a)\ (s-b)\ (s-c)/s)\ .$$

Angle θ obeys the relation

$$(10)\quad \theta = 2 \cdot \arctan\ (v/(s-a)).$$ (Burington, 1948, pg. 20).

The co-ordinates of the two possible points for C_z are

$$(11a)\quad X_z = X_x + \cos(\theta) \cdot b$$

and

$$= X_x + \cos(\theta) \cdot L_{xz}$$

$$(11b)\quad Y_z = Y_x \pm \sin(\theta) \cdot b$$
$$= Y_x \pm \sin(\theta) \cdot L_{xz}\ .$$

The program in Appendix 1 has an option which locates all circles relative to $C_X = (0,0)$ using the scale R = 1, or, as an option, the user may specify the desired scale and origin. The program then locates the ballentine on the user's co-ordinate system.

ERIC Full Text Provided by ERIC

Legends

Figure 1: A ballantine representing the correlations between an auditory detection task, a visual detection task, and a test of arithmetic skill.

Figure 2: A correlation indicated by an overlap between two circles. For the representation to be exact the proportion of the area of each circle that falls in region A should be equal to r^2.

Figure 3: The geometric relations used to construct an appropriate ballantine. Angle α is implicitly defined by r^2. Angle α, in turn, determines the length of line L_{xy}.

Figure 4: The three lines between the centers of the circles define one of two possible triangles, with Circle Z either above or below line L_{xy}. By solving for the interior angle θ at the center of circle X, and given L_{xz}, the position of Circle Z is determined relative to X and the position of Circle X.

References

Burington, R.S. (1948) Handbook of mathematical tables and formulas. Sandusky, OH: Handbook Publishers, Inc.

Cohen, J. and Cohen, P. (1975) Applied multiple regressional and correlation analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum Associates.

Acknowledgement Note

The preparation of this paper was supported in part by Contract N00014-84-K-5553 between the Office of Naval Research and the University of Washington (Earl Hunt, Principal Investigator). The opinions expressed are entirely the responsibility of the author. Colene McKee's assistance in preparing the graphs is gratefully acknowledged.

Appendix 1

```
program ballantine(input,output);
( Locates circles so that overlap is r^2 of area of each circle )

const pi = 3.14159265;

var cx, cy, radius :real;    com: char;

function a2srqq(consts p1,p2:real4): real4; extern;
     ( IBM arctan function )

function converge(r:real):real;
     (computes value of angle theta , and then
uses theta to compute the distance between circle centers.
     Input is r^2.
Output is distance between centers, assuming radius of 1 )

var high,low, alpha, old, delta, q, z : real;

begin ( converge )
if r = 1.0 then converge := 0.0 ( identical circles ) else
if r <= 0.0 then converge := 2.0 (no intersection ) else
```


ERIC Full Text Provided by ERIC

```
begin  { intersecting circles. compute overlap }
  low := 0; high := pi; { angle theta from 0 to pi in radians }
  q := r * pi;  { sector area of overlap }
  old := 0; alpha := pi/2.0; { 90 degrees, initial guess for angle }

  repeat { converge loop }
    z := alpha -sin(alpha);
    if z = q then delta := 0.0  {exact match} else
    {compute adjustment }
      begin
      old := alpha;
      if z > q then { decrease alpha }
        begin
        alpha := alpha - (alpha-low)/2.0;
        high := old;
        end
      else { increase alpha }
        begin
        alpha := alpha + (high-alpha)/2.0;
        low := old;
        end;
     delta := abs(old-alpha);  {size of adjustment }
     end; { of adjustment }
   until delta <0.0001; { converge to thousandth of a radian }
   { compute the distance between circles }
```



```
    alpha := alpha/2.0;
    converge := 2.0 * cos(alpha);
   end; { overlap computed }
  end;  { converge function }
  procedure graphpars(var cx,cy,radius :real);
  {Computes the scale and translation factors for a real graph }

  begin
    writeln ('Your graph is assumed to have 0,0 at the lower left');
    writeln ('enter maximum value of x and y as integers ');
    readln(cx,cy);
    cx := cx/2.0; cy := cy/2.0;
    if cx < cy then radius :=0.90 * cx/2.0
       else radius := 0.9 * cy/2.0;
  end; { Graphpars }

  Procedure twocircles(cx,cy,radius:real);

  var r,c,l,z,x :real;

  begin { twocircles }

          writeln ('What is the value of the correlation ?');
          readln(r); r := r * r;
          l := converge(r) * radius ;
```

```
     writeln('Distance between circles is ',l:10:4);
     x := cx-l/2.0; z := cx + l/2.0;
     writeln( 'Circle X at point ',x:7:2,' ',cy:7:2);
     writeln( 'Circle Y at point ',z:7:2,' ',cy:7:2);
     writeln(' Radius = ',radius:10:4);
end; { twocircles }

Procedure threecircles(cx,cy,radius:real);

const x = 1; y = 2; z = 3; { used for names of circle }
var  rxy, rxz, ryz, lxy, lxz, lyz :real ; { Same names as in paper }
     cc : array [1..3,1..2] of real; { centers of circles }
     a,b, c, s, theta, v : real; { Auxilary variables named in paper }
     xx, dx, dy : real;  { scratch variables for computing }

begin { procedure threecircles }

{get needed values }
     writeln ('Values of correlations rxy, rxz, ryz (real ) ');
     readln (rxy,rxz,ryz);
     rxy := rxy * rxy; rxz := rxz * rxz; ryz := ryz * ryz;
{ calculate intercircle distances }
       lxy := converge(rxy);
       lxz := converge(rxz);
       lyz := converge(ryz);
```



```
{ convert to auxilary notation to conform to the text }
       a := lyz; b := lxy; c := lxz;
       s := (a + b + c)/2.0;
       v := (s-a) * (s-b) * (s-c) / s;
       v := sqrt(v); xx := s-a;
{ calculate value of interior angle theta at center of circle x and then
  determine the distance center of z falls below the x-y centerline. }
       theta := 2 * a2srqq(v,xx); { IBM terminology for arctan }
       dy := sin(theta) * lxz; dx := cos(theta) * lxz;
{determine center points,converting to actual graph }
       cc[x,x] := cx - (lxy/2.0) * radius;
     {x-y symmetric re vertical axis}
       cc[x,y] := cy + (dy/2.0) * radius;
      { x-z symmetric re horizontal axis}
       cc[y,x] := cx + (lxy/2.0) * radius;
       cc[y,y] := cc[x,y];
       cc[z,x] := cc[x,x] + dx * radius;
       cc[z,y] := cc[x,y] - dy * radius;
{print results }
       writeln ('Ballantine for rxy = ',sqrt(rxy):5:3,
              ' rxz= ',sqrt(rxz):5:3,' ryz ',sqrt(ryz):5:3);
       writeln;
       writeln ('circle    X      Y');
       writeln ('  X    ',cc[x,x]:7:2,' ',cc[x,y]:7:2);
       writeln ('  Y    ',cc[y,x]:7:2,' ',cc[y,y]:7:2);
```

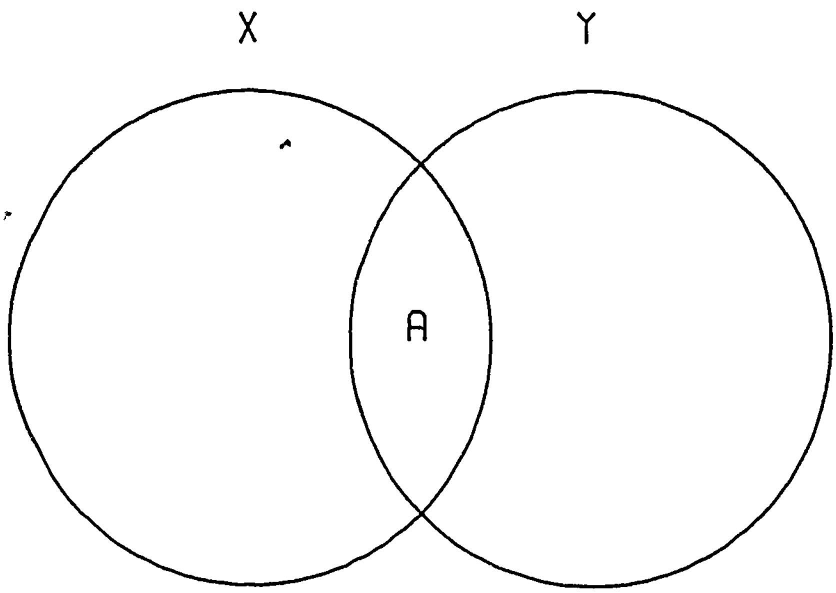
```
        writeln ('  Z   ',cc[z,x]:7:2,' ',cc[z,y]:7:2);
        writeln;
        writeln (' All radii = ',radius:7:1);

end; { Of threecircle procedure }

begin { main program }
  write (' Is a real (r) or abstract (a) graph to be positioned ' );
  readln(com);
  if com = 'r' then graphpars(cx,cy,radius) else
     begin
     writeln ('Abstract graph centered at 0,0 with radius = 1 ');
     cx := 0; cy := 0; radius := 1.0;
     end;
  write(' Is a two (2) or three (3) variable problem to be computed? ');
  readln(com);
  if com = '2' then twocircles(cx,cy,radius) else
  if com = '3' then threecircles(cx,cy,radius)
       else writeln ('undefined problem');
end. (Of main program )
```

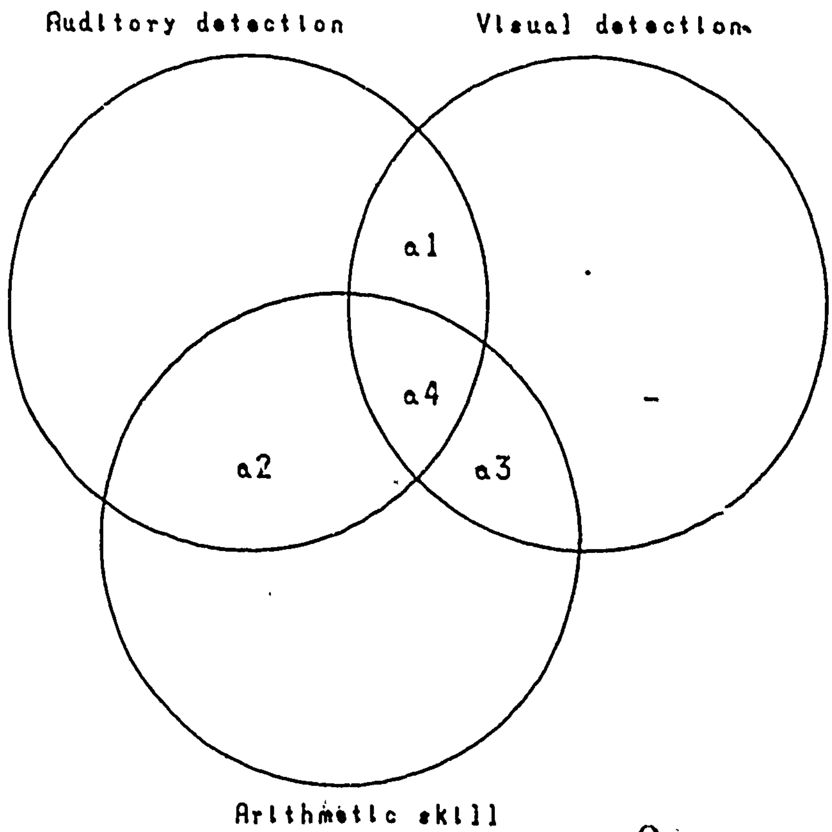

Table 1

r	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	2.00	1.92	1.87	1.83	1.79	1.76	1.72	1.69	1.67	1.64
0.10	1.61	1.58	1.56	1.53	1.51	1.49	1.46	1.44	1.42	1.40
0.20	1.37	1.35	1.33	1.31	1.29	1.27	1.25	1.23	1.21	1.19
0.30	1.17	1.15	1.13	1.11	1.09	1.08	1.06	1.04	1.02	1.00
0.40	0.98	0.97	0.95	0.93	0.91	0.89	0.88	0.86	0.84	0.83
0.50	0.81	0.79	0.77	0.76	0.74	0.72	0.71	0.69	0.67	0.66
0.60	0.64	0.62	0.61	0.59	0.57	0.56	0.54	0.52	0.51	0.49
0.70	0.48	0.46	0.44	0.43	0.41	0.40	0.38	0.36	0.35	0.33
0.80	0.32	0.30	0.28	0.27	0.25	0.24	0.22	0.20	0.19	0.17
0.90	0.16	0.14	0.13	0.11	0.09	0.08	0.06	0.05	0.03	0.02

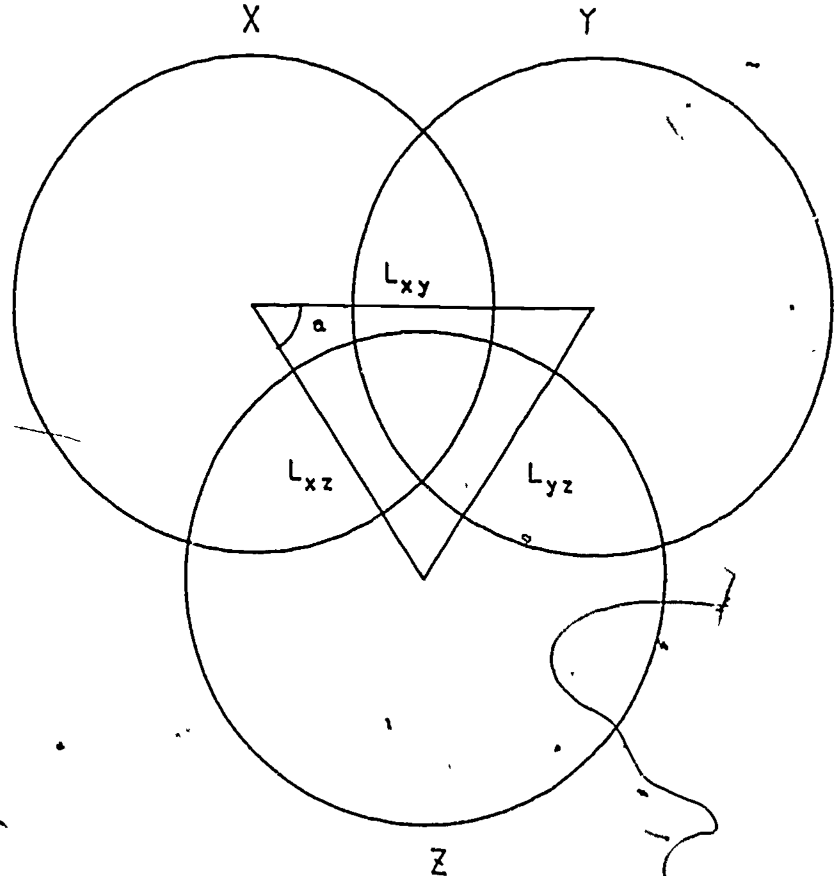
Distance between circles of radius one as a function of the correlation (r) between the variables represented by the circles

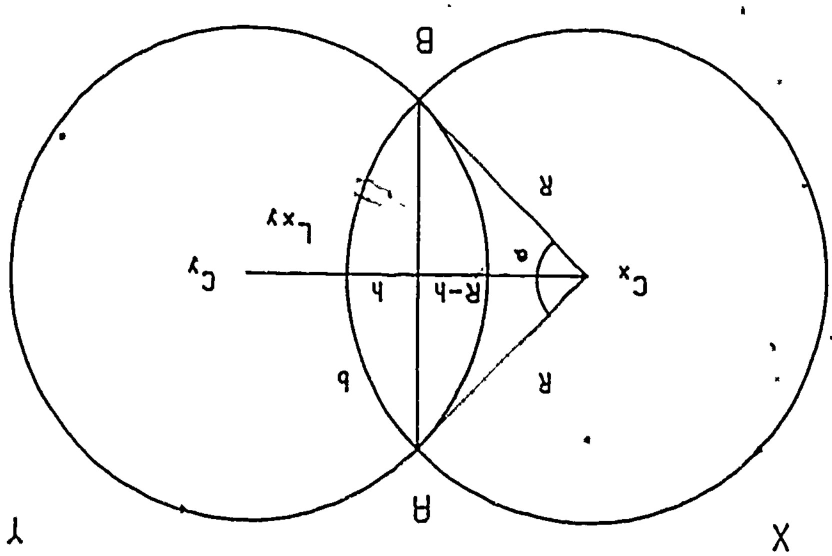
ERIC


X
Y
A


Auditory detection
Visual detection
a1
a4
a2
a3
Arithmetic skill

ERIC


X
Y
L_{xy}
a
L_{xz}
L_{yz}
Z


B
R
L_{xy}
a
C_y
C_x
h
R-h
b
R
A
Y
X